

第一篇论文初步阅读报告

该实验采用基于 CNN 的视频分割方法，每个序列进行 3-4 次人工分割训练。

该方法结合了两个算法。

- 一个是根据表面位移来细化参数化的人体模型。
- 另一个是一种新的方法，即在一个共同的参照系中对动态人体轮廓锥进行变形和融合。融合锥融合了视频中包含的形状信息，让我们能够经过优化得到详细的模型形状。该算法不仅能捕捉到人脸的几何形状和外观，还能自动地获得运动学骨架的人体模型，从而实现与姿势有关的近似的表面变形。

1. SMPL 模型改进：增加补偿项

SMPL 模型输入是 10 种身形参数 (β , 包含人的高矮胖瘦和头身比) 和 72 种姿态参数 (θ , 包含运动位姿以及 24 个关节相对角度)。输出是 $N=6890$ 个顶点的 triangulated mesh。

T: 身形和姿态的变形。

W: 一个混合蒙皮 (blend-skinning) 线性方程，其包含：

- 骨骼关节 $J(\beta)$
- 与姿态有关的变形 $B_p(\theta)$
- 与身形有关的变形 $B_s(\beta)$

因为 SMPL 是裸的人体模型，而该实验最后的目标并不是一个裸露的人的模型，而是包含衣服，头发甚至纹理，所以在 T 的表达式中加了一个补偿项 D，**该项会在第二步得到优化**：

$$T(\theta, \beta, D) = T_\mu + B_s(\beta) + B_p(\theta) + D$$

2. 姿态重建(Pose Reconstruction)

计算每个帧的 3D 姿势。输出是成序列的 F 帧的一个 pose 集合。

为了避免尺度模糊 (scale ambiguity)，每次计算时只考虑 5 帧，即一个身形的五个姿势。

首先，以 SMPL 模型的身形参数 β_0 为固定的初始化值，为了更精确，计算时引入轮廓项。为进一步获得极小值，在 Gaussian 金字塔 G 的四个层次上进行了优化。

$$E_{\text{silh}}(\theta) = G(w_0 I_{rn}(\theta)C + w_i(1 - I_{rn}(\theta))\bar{C}) \quad (4)$$

接下来使用二维关节检测和以 SMPL 中单模态的 A 姿态先验的方法来计算其姿势。用 SMPL 中 pose 参数的来训练这个 prior。即进行一系列的迭代，第一帧的初始化值就是 SMPL 中的 A 姿态，而在预估新的一帧中的姿态时，初始化值是上一帧的最终预估值。如果误差过大则将 pose 参数重新初始化为 0。除此以外，该试验方法的创新点是在计算时引入轮廓项，并在 Gaussian 金字塔的四个层次上进行了优化。最终得到的结果是一段时间内的一系列姿态。

3. 共识身形估计(Consensus Shape Estimation)

对于上一步得到的一系列姿势，从中分批给定一组估计的姿势，可以用它们联合优化单个精确的身形来匹配所有原始的姿势。因为在每一帧人的轮廓都会给身形带来新的限制，具体来说，从相机到轮廓点的光线集合定义了一个约束锥。该实验就是将顶点 v 的变形函数（SMPL 函数）进行逆变换运用到轮廓锥的每一条光线 r 中。对于每一条射线 r ，都能找到它最近的 3D 模型点。对模型点对应的射线 r 进行逆变换。这样就能有效地消除轮廓锥，并对标准 T 形的身形设置约束，每一批次训练 120 个轮廓，经过计算得到的光线形成了一个规范的 T 型模型。这一步骤叫做 **unposing**，它使得身体形状得到有效优化并独立于姿态。

具体地，该实验构建了一个能量式，即 1 个数据能量和 3 个正则化能量的线性组合。

- 第一项：轮廓项，测量边界点和轮廓线之间的距离
- 第二项：拉普拉斯项（用来保证形变的平滑性）
- 第三项：身体模型项（进行偏差处罚，即误差消减）
- 第四项：对称项（约束用来保证形体的对称性）。通过身形参数 β 以及补偿项 D 的改变可以使得整个式子的值达到最小从而得到精确的身形。

$$E_{\text{cons}} = E_{\text{data}} + w_{\text{lp}} E_{\text{lp}} + w_{\text{var}} E_{\text{var}} + w_{\text{sym}} E_{\text{sym}} \quad (7)$$

4. 帧细化和纹理的映射生成(frame refinement and texture map generation)

在计算了给定序列帧的全局形状后，接下来的目标是捕获时间变化。

将第二步的能量式修改后按照顺序来处理一系列帧。具体是用前一帧计算所得来。

初始化优化，并与相邻帧进行正则化优化。为了创建纹理，将估计的规范模型返回到每个帧，将图像颜色向后投影到所有可见顶点，最后通过计算所有视图中最正交的纹理的中值生成纹理图像。